

SIMPOETS, CEFET-GO, 86-103, 2008

Estudos de Confiabilidade na Modelagem de Sistemas

Victor Freitas de Azerêdo Barros

Bacharelando em Informática, Bolsista CNPq, CEFET-GO UNED Inhumas
vfbarros@msn.com

José Elmo de Menezes

Doutor em Informática, Orientador, CEFET-GO UNED Inhumas e UCG-GO
elmo@inhumas.cefetgo.br

Resumo: Atualmente, as empresas buscam maior confiabilidade e qualidade em seus produtos por causa da grande concorrência de mercado. Além disso, a diminuição dos gastos com garantia e a possibilidade de extensão da mesma é um grande diferencial de mercado, e pode ser obtido com a realização de estudos de confiabilidade dos sistemas e equipamentos. A confiabilidade é uma grandeza mensurável a partir da análise estatística da ocorrência de falha. Falha é o fenômeno que compromete o desempenho ou impossibilita a operação de um item. Assim, um conceito de confiabilidade é a “probabilidade de um sistema ou item desempenhar satisfatoriamente a função requerida, sob condições de operações estabelecidas, por um período de tempo pré-determinado”, Lewis (1996).

Palavras-chave: Sobrevivência, Falha, Confiabilidade.

1. Introdução

O estudo de modelagem estatística de tempos de vida de sistemas está baseado em conceitos básicos de confiabilidade. Assim, um conceito de confiabilidade é a “*probabilidade de um sistema ou item desempenhar satisfatoriamente a função requerida, sob condições de operações estabelecidas, por um período de tempo pré-determinado*” (Lewis, 1996).

O Cliente quando adquire um produto, está aceitando que este poderá falhar em algum instante no futuro. Este enfoque é frequentemente acoplado com uma garantia, que pode estar em lei ou não, de forma que ele possa reivindicar de falhas que possam ocorrer dentro de um período de tempo específico (O’CONNOR in VOLLERTT JUNIOR, 1996).

Fabricantes estão se conscientizando de que a confiabilidade de seus produtos acarreta um menor custo de reparo durante os prazos de vigência das garantias, assim como a perda da imagem da empresa através da insatisfação dos clientes. Por esta razão e devido ao mercado competidor, que está oferecendo aos clientes prazos de garantia cada vez maiores, fabricantes estão estudando e aplicando os métodos da engenharia de confiabilidade a fim de proporcionar maior credibilidade de seus produtos e também para poder prever o tempo permitido e os gastos envolvidos com as garantias (VOLLERTT JUNIOR, 1996).

Quando um equipamento que até então estava funcionando normalmente falhe e que esse estado muda à medida que o tempo evolui com certa probabilidade e que, essa probabilidade de transição depende apenas do estado em que se encontra e do estado a seguir (BOLDRINI e COSTA e FIGUEIREDO e WETZLER, 1980), então assumimos que essa abordagem para análise do processo de trocas de estado é denominada de cadeia de Markov (HENLEY e KUMAMOTO, 1981). A cadeia de Markov é um processo estocástico em que o próximo estado depende apenas do seu estado atual, ou seja, as informações de estados passados são irrelevantes e o tempo que o processo está no estado atual também é irrelevante (BRESSAN, 2002).

Atualmente são utilizados vários programas estatísticos para a modelagem de sistemas, sendo que no presente estudo foi adotado o programa estatístico R. Os comandos utilizados no R auxiliam na obtenção dos modelos de sobrevivência e fracasso, nas estimativas de Kaplan-Meier e na escolha do modelo probabilístico na análise de dados de sobrevivência. (REIS e JUNIOR, 2007). O estimador não-paramétrico de Kaplan-Meier, proposto por Kaplan e Meier (1958) para estimar a função de sobrevivência, é também chamado de estimador limite-produto (COLOSSIMO e GIOLO, 2006).

O comportamento do tempo de sobrevivência pode estar associado a várias causas e pode ser de difícil representação matemática, sendo assim a forma mais simples e eficiente de selecionar o “melhor” modelo a ser usado em um conjunto de dados é por meio de técnicas gráficas (LOUZADA NETO e MAZUCHELI e ACHCAR, 2002). Existe uma série de modelos probabilísticos utilizados em análise de dados de sobrevivência. Neste trabalho foram utilizados os modelos exponencial, de Weibull e log-normal.

A distribuição exponencial apresenta um único parâmetro e é a única que se caracteriza por ter uma função de taxa de falha constante. A distribuição de Weibull (1939) apresenta uma grande variedade de formas, todas com uma propriedade básica: a sua função de taxa de falha é monótona, isto é, ela é crescente, decrescente ou constante. A distribuição log-normal pode ser analisada segundo uma distribuição normal, desde que, se considere o logaritmo dos dados em vez de valores originais (COLOSSIMO e GIOLO, 2006).

O trabalho teve como objetivo um estudo aprofundado na teoria de Probabilidade e Estatística, trabalhando com simulações e modelos, utilizando o software R.

2. Materiais e Métodos

O estudo foi baseado na cadeia de Markov através da análise de trocas de estados em sistemas binários de falha e reparo. De uma amostra de 10 componentes, todos começando do seu estado normal ($t=0$), verificou-se o tempo de falha e o tempo de reparo. No momento em que o componente falha ele instantaneamente entra no processo de reparo e, após o tempo de reparo, ele instantaneamente volta ao seu estado normal, assim sucessivamente.

Com os tempos de falha e de reparo obtêm-se uma tabela e com ela foi calculado:

4. A confiabilidade, $R(t)$, que pode ser representada como a probabilidade de sobrevivência do equipamento até um determinado tempo t ;
5. A probabilidade de ocorrência de falha, $F(t)$, que avalia a possibilidade de um produto falhar num determinado período de tempo t ;
6. A função densidade de probabilidade, $f(t)$, que é um histograma represando o número de falhas num período de tempo t ;
7. A taxa de falha, $r(t)$, que é a probabilidade de falha por unidade de tempo t ;
8. A disponibilidade, $A(t)$, que pode ser representada como a probabilidade do componente estar no seu estado normal no tempo t ;
9. A intensidade de Falha Incondicional, $w(t)$, que é a probabilidade do componente falhar por unidade de tempo t ;
10. A intensidade de Falha Condicional, $\lambda(t)$, que é a proporção da população no seu estado normal que se espera a falhar por unidade de tempo t . Grandes valores de λ significam que os componentes estão prestes a falhar.

Diante dos dados citados acima, foi elaborada uma segunda tabela com os dados de mortalidade obtidos a partir do processo de reparo e falha. A partir destes momentos foi possível traçar, utilizando-se dos dados da função $R(t)$ e $F(t)$, as curvas de sobrevivência e fracasso. Estas curvas foram obtidas através do programa estatístico R por meio de comandos específicos.

A partir dos tempos de reparo foi utilizado o estimador de Kaplan-Meier, construindo um gráfico da função de sobrevivência estimada por Kaplan-Meier, através do programa R. Através deste gráfico foi feito um estudo com modelos probabilísticos para ajustar os modelos apresentados (exponencial, de Weibull e log-normal) e, com base na comparação entre valores estimados e observados, decidir qual deles “melhor” explica os dados amostrais, ou seja, relacionar qual modelo que mais se aproxima à curva de sobrevivência.

3. Resultados e Discussões

Consideremos o exemplo extraído de Henley e Kumamoto (1981), em que o seu estudo foi baseado em conceitos de confiabilidade e modelos estatísticos. As transições de repetições de falha e reparo em um sistema binário estão descritas na Figura 1.

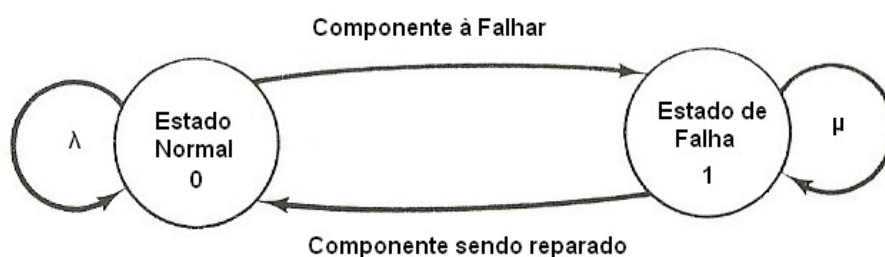


Figura 1. Diagrama dos Estados de Transições dos Componentes

Com o diagrama, vimos que um componente que até então estava funcionando corretamente pode falhar, e que seu estado (normal ou falha) muda à medida que o tempo evolui. Essas experiências de falha são chamadas de estado de transição. O estado de falha continua sempre se não for reparado e o componente a ser reparado permanece no período de falha por um tempo e, em seguida, sofre uma transição para o estado normal quando o reparo for concluído.

A transição para o estado normal (0) é chamada de reparação, e a transição para o estado de falha (1) é chamada de fracasso. Assumimos que, os reparos restauram o componente a uma condição tão boa como nova e que todo o ciclo, portanto, consiste de constantes repetições de processos de reparação/fracasso e fracasso/reparação.

Henley e Kumamoto, em seu livro *Reliability Engineering and Risk Assessment* (1981), descreve de uma maneira simplificada uma forma de produzir um histórico de componentes como mostrado na Figura 2.

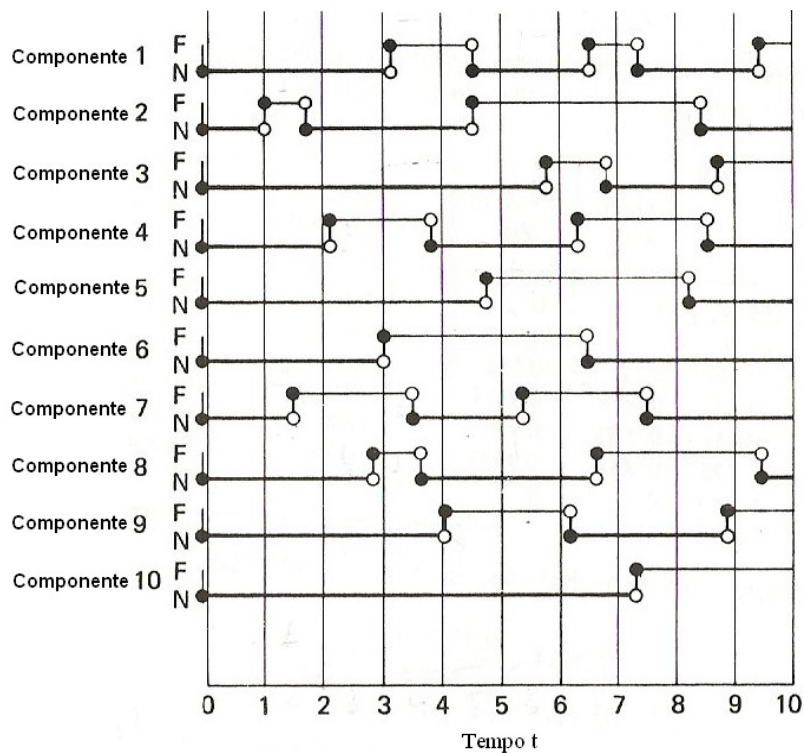


Figura 2. Histórico dos Estados dos Componentes

Essas características desses componentes com repetidos processos de reparação/fracasso podem ser obtidos mediante a ponderação da componente como uma amostra de uma população semelhante sujeito a repetições idênticas. Todas as amostras começam em seu estado normal no momento nulo; ou seja, cada componente é tão bom como um novo em $t = 0$.

Os tempos obtidos do histórico dos estados dos componentes foram transcritos para a tabela 1, com o tempo que leva para o estado normal falhar e o tempo em que ele é reparado.

Tabela 1. Tempo de Reparo e Falha dos componentes

Componente	Reparo	Falha
01	0.0	3.1
	4.5	6.6
	7.4	9.5
02	0.0	1.05
	1.7	4.5
03	0.0	5.8
	6.8	8.8
04	0.0	2.1
	3.8	6.4
05	0.0	4.8
06	0.0	3.0
07	0.0	1.4
	3.5	5.4
08	0.0	2.85
	3.65	6.7
09	0.0	4.1
	6.2	8.95
10	0.0	7.35

Com esses dados, podem-se obter os dados de mortalidade, a partir dos momentos de fracasso, podendo se calcular a confiabilidade, a probabilidade de ocorrência de falha, a função densidade de probabilidade e a taxa de falha, conforme a tabela 2.

Tabela 2. Dados de Mortalidade e do Processo de Reparação/Fracasso

t		R(t)	F(t)	f(t)	r(t)
0	[0,1)	1.0000	0.0000	0.0000	0.0000
1	[1,2)	1.0000	0.0000	0.1111	0.1111
2	[2,3)	0.8889	0.1111	0.1111	0.1250
3	[3,4)	0.7778	0.2222	0.1111	0.1428
4	[4,5)	0.6667	0.3333	0.1667	0.2500
5	[5,6)	0.5000	0.5000	0.1111	0.2222
6	[6,7)	0.3889	0.6111	0.1667	0.4286
7	[7,8)	0.2222	0.7778	0.0556	0.2502
8	[8,9)	0.1667	0.8333	0.1111	0.6665
9	[9,10)	0.0556	0.9444	0.0556	1.0000

Para a obtenção dos dados de confiabilidade, representado por $R(t)$, calcula-se a probabilidade de um equipamento não falhar até um certo tempo t . Em termos probabilísticos, isto é escrito como:

$$R(t) = P(T \geq t) \quad (1)$$

Em consequência, para a obtenção dos dados de $F(t)$, que é a probabilidade de ocorrência de falha, calcula-se a possibilidade de um equipamento falhar no período de tempo t , expressada como:

$$F(t) = 1 - R(t). \quad (2)$$

Verifica-se que quanto maior o tempo de operação do equipamento, maior a probabilidade da ocorrência de falha, portanto, para tempos muito longos, a probabilidade é de que 100% dos equipamentos estejam no seu estado de falha.

Foi calculada a função densidade de probabilidade, $f(t)$, que nada mais é que um histograma, representando o número de falhas num período de tempo, que é dada como:

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t)}{\Delta t} \quad (3)$$

A taxa de falha, $r(t)$, é a probabilidade de morte por unidade de tempo t , sendo definida como o limite da probabilidade de um indivíduo falecer no intervalo de tempo t , dado que o mesmo tenha sobrevivido até o tempo t , sendo expresso por:

$$r(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t} \quad (4)$$

Esta função também pode ser definida por meio da expressão:

$$r(t) = \frac{f(t)}{R(t)} \quad (5)$$

Henley e Kumamoto dizem que a taxa de falha também pode ser representada por uma curva que possui a forma de uma banheira, que é caracterizada por uma elevada taxa relativamente precoce, chamada de região operacional, seguida de uma taxa relativamente constante, chamada de região constante ou vida útil que é o principal período que ocorrem falhas aleatórias e, em seguida, por uma taxa de falha crescente, chamada de região de desgaste.

Também foi calculada a disponibilidade, $A(t)$, que pode ser representada como a probabilidade do componente estar no seu estado normal no tempo t , ou seja, o número de componentes normais no tempo t dividido pela amostra total.

A Intensidade de Falha Incondicional, $w(t)$, que é a probabilidade do componente falhar por unidade de tempo no tempo t , ou seja, o número de componentes no estado de falha no tempo t dividido pela amostra total.

A Intensidade da Falha Condicional, $\lambda(t)$, que é a proporção da população no seu estado normal que se espera a falhar por unidade de tempo, ou seja, o número

de componentes no estado de falha dividido pela população no seu estado normal. Grandes valores de $\lambda(t)$ significam que os componentes estarão prestes a falhar.

Tendo apenas como exemplo para cálculo, usando a Figura 2, ao se calcular a disponibilidade, a falha incondicional e a intensidade de falha condicional, no período de tempo 5, ou seja $[5,6)$, temos como resultado:

$$A(5) = \frac{6}{10} = 0.6 \quad (6)$$

$$w(5) = \frac{2}{10} = 0.2 \quad (7)$$

$$\lambda(5) = \frac{2}{6} = 0,333 \dots \quad (8)$$

Outra abordagem para a análise de todo o processo para o caso de falha constante e taxas de reparação é a análise de Markov. Sendo x a variável indicadora (0,1). A definição da Intensidade da Falha Condicional, $\lambda(t)$, pode ser usada para fornecer a probabilidade de falha em $t + dt$, dado que o componente esteja trabalhando no tempo t .

$$P(1 | 1) \equiv Pr[x(t + dt) = 1 | x(t) = 1] = 1 - \lambda dt \quad (9)$$

$$P(0 | 1) \equiv Pr[x(t + dt) = 0 | x(t) = 1] = \lambda dt \quad (10)$$

$$P(0 | 0) \equiv Pr[x(t + dt) = 0 | x(t) = 0] = 1 - \mu dt \quad (11)$$

$$P(1 | 0) \equiv Pr[x(t + dt) = 1 | x(t) = 0] = \lambda dt \quad (12)$$

A curva, $R(t) \times t$, é uma função de sobrevivência, e, a curva, $F(t) \times t$, é uma função de fracasso. Ambas as curvas são representadas pela figura 3 abaixo.

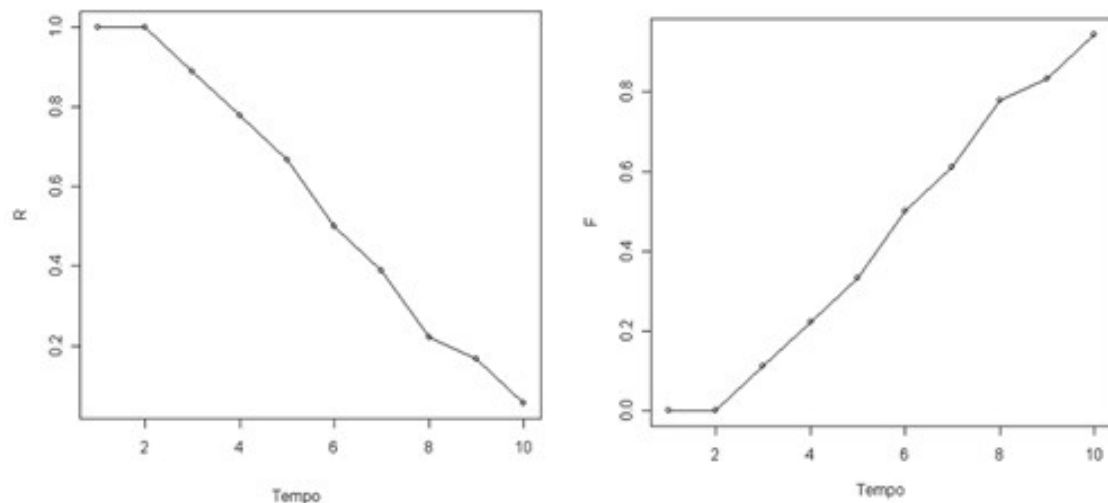


Figura 3. Função de Sobrevivência e Função de Fracasso

A curva $R(t|xt)$, é uma função de sobrevivência que representa tanto a probabilidade de sobrevivência de um indivíduo de idade t como também a proporção da população que se espera que venha a sobreviver a qualquer idade t , enquanto que a curva $F(t|xt)$, é uma função de fracasso que representa tanto a probabilidade de morte antes da idade t como também a proporção da população que está prevista a morrer antes da idade t .

Colossimo e Giolo em seu livro *Análise de Sobrevivência Aplicada* (2006), usa um exemplo de aplicação de um conjunto de técnicas estatísticas para dados da área da saúde. Aqui usaremos a mesma seqüência cronológica, estimando a curva de sobrevivência pelo modelo de Kaplan-Meier e pelos modelos exponencial, de Weibull e log-normal, porém Colossimo e Giolo calculou com os tempos de reincidência de um grupo de pacientes com câncer de bexiga submetidos a um procedimento cirúrgico feito por laser, e aqui foi calculado com os tempos de falha de uma amostra de 10 componentes. A figura 4 mostra o gráfico da função de sobrevivência estimada por meio do Kaplan-Meier, a partir dos tempos de falha.

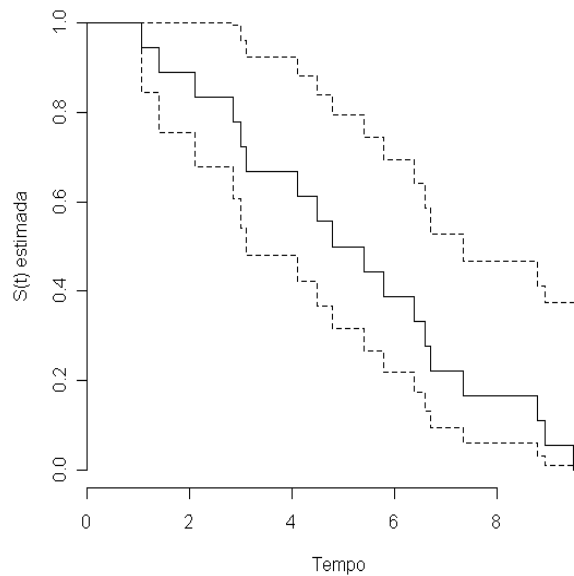


Figura 4. Gráfico da estimativa de Kaplan-Meier

Com as expressões das estimativas da função de sobrevivência para os modelos exponencial, de Weibull e log-normal obteve os seguintes dados, respectivamente,

$$\hat{S}(t)_e = \exp \left\{ - \left(\frac{t}{5,13} \right) \right\} \quad (13)$$

$$\hat{S}(t)_w = \exp \left\{ - \left(\frac{t}{5,80} \right)^{2,16} \right\} \quad (14)$$

$$\hat{S}(t)_{ln} = \phi \left[- \frac{(\log(t) - 1,48)}{0,62} \right] \quad (15)$$

Esses valores apresentados nas expressões são as estimativas de cada um dos modelos.

A tabela 3 mostra as estimativas das funções de sobrevivência para os tempos de falha usando os modelos exponencial, de Weibull e log-normal e também o Kaplan-Meier.

Tabela 3. Estimativas de sobrevivência

t	K-M	Exp	Weibull	Log-normal
1.05	0.9994	0.8150	0.9455	6.5918e ⁻³
1.40	0.8889	0.7613	0.9549	1.6092e ⁻³
2.10	0.8333	0.6643	0.8949	1.5628e ⁻⁴
2.85	0.7778	0.5740	0.8065	2.0594e ⁻⁵
3.00	0.7222	0.5574	0.7864	1.4319e ⁻⁵
3.10	0.6667	0.5467	0.7726	1.1312e ⁻⁵
4.10	0.6111	0.4499	0.6235	1.3566e ⁻⁶
4.50	0.5556	0.4162	0.5611	6.4065e ⁻⁷
4.80	0.5000	0.3926	0.5146	3.7593e ⁻⁷
5.40	0.4444	0.3493	0.4243	1.3826e ⁻⁷
5.80	0.3889	0.3231	0.3677	7.4071e ⁻⁸
6.40	0.3333	0.2874	0.2899	3.0687e ⁻⁸
6.60	0.2778	0.2765	0.2663	2.3179e ⁻⁸
6.70	0.2222	0.2711	0.2549	2.0192e ⁻⁸
7.35	0.1667	0.2389	0.1882	8.5241e ⁻⁹
8.80	0.1111	0.1801	0.0849	1.4961e ⁻⁹
8.95	0.0556	0.1749	0.0775	1.2652e ⁻⁹
9.50	0.0000	0.1571	0.5447	6.9621e ⁻¹⁰

Para a escolha de um dos modelos, construíram-se os gráficos das estimativas de sobrevivência obtidas pelo método de Kaplan-Meier *versus* as estimativas de sobrevivência obtidas a partir dos modelos de exponencial, de Weibull e log-normal, respectivamente, conforme a figura 5.

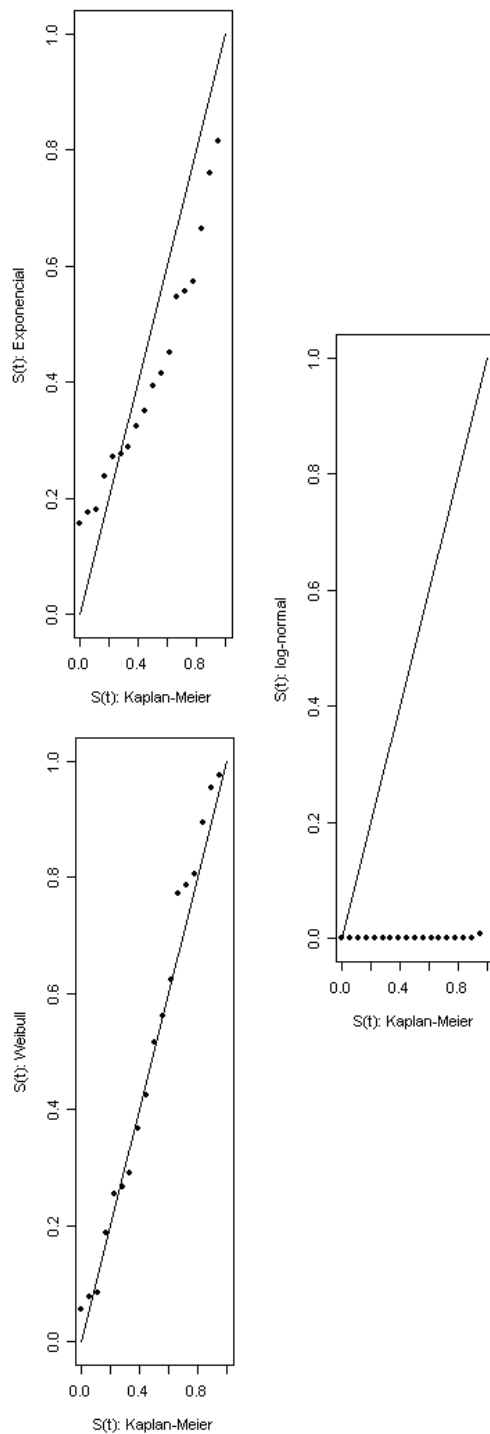


Figura 5. Gráficos das sobrevivências estimadas por Kaplan-Meier *versus* as sobrevivências estimadas pelos modelos exponencial, de Weibull e log-normal

A partir dos gráficos é possível observar que o modelo log-normal parece não ser adequado para esses dados, pois a curva se apresenta um tanto afastada da reta $y=x$. Por outro lado, os modelos exponencial e de Weibull acompanham mais de perto a reta $y=x$, indicando ser um dos modelos, possivelmente, adequado para os dados sob estudo.

Na tentativa de confirmar os resultados já obtidos, foram construídos os gráficos linearizados para os modelos exponencial, de Weibull e log-normal, respectivamente, conforme a figura 6.

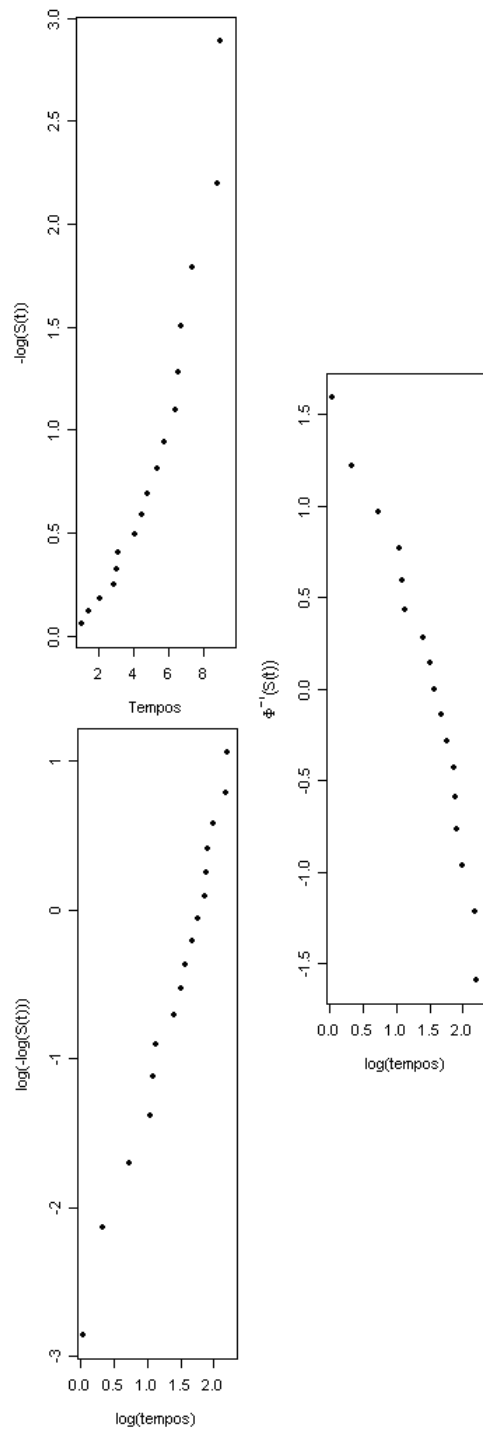


Figura 6. Gráficos linearizados para os modelos exponencial, de Weibull e log-normal

Os gráficos para os modelos exponencial e de Weibull apresentados acima não mostram afastamentos marcantes de uma reta. Já para o modelo log-normal, observa-se um certo desvio de reta. Esses gráficos confirmam os resultados observados e indicam os modelos de exponencial e de Weibull para serem usados na análise dos dados.

Os dois modelos indicados pelos procedimentos gráficos apresentaram resultados similares e igualmente bons. Para que assim confirme que ambas, exponencial e de Weibull, apresentam ajustes satisfatórios, fazemos um novo gráfico agora com as curvas de sobrevivência estimadas por meio do ajuste de ambos os modelos *versus* a curva de sobrevivência estimada por Kaplan-Meier que podem ser observadas nas figuras 7 e 8.

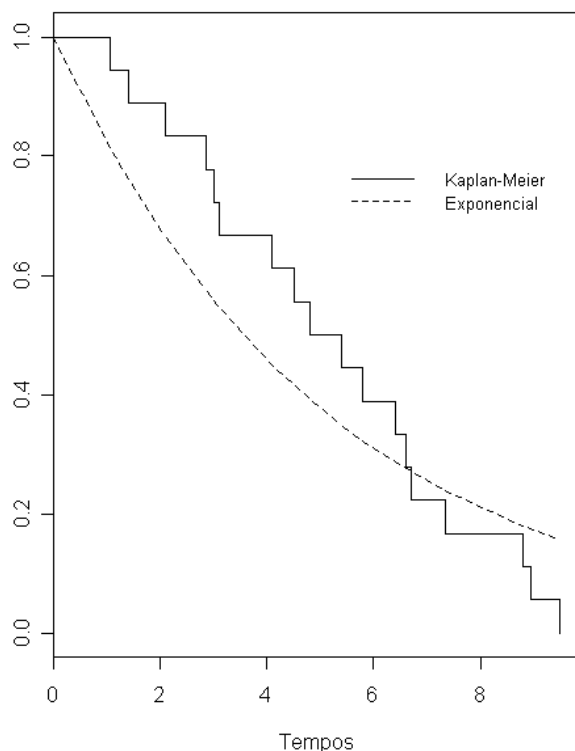


Figura 7. Gráfico da curva de sobrevivência estimada pelo modelo exponencial *versus* a curva de sobrevivência de Kaplan-Meier

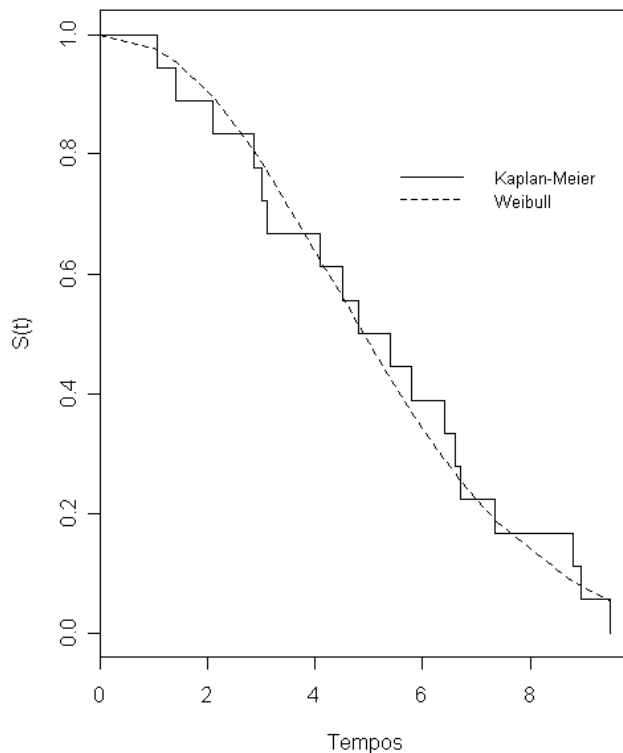


Figura 8. Gráfico da curva de sobrevivência estimada pelo modelo de Weibull *versus* a curva de sobrevivência de Kaplan-Meier

Por fim, podemos confirmar que o modelo de Weibull, é o modelo mais adequado, pois, é o modelo que cujos pontos da função de sobrevivência estão mais próximos dos valores obtidos pelo estimador de Kaplan-Meier.

Colossimo e Giolo (2006) encontraram resultados parecidos ao deste estudo, ao analisar o tempo de reincidência de um grupo de pacientes com câncer de bexiga submetidos a um procedimento cirúrgico feito por laser, porém, para eles os modelos que apresentaram os melhores ajustes foi o modelo de Weibull e log-normal.

4. Conclusões

Com a utilização da cadeia de Markov foi possível analisar e calcular os processos de trocas de estado de um sistema binário, portanto, esta é válida para modelar estatisticamente os tempos de vida de sistemas. Verificou-se também que o programa R é confiável para o desenvolvimento dos modelos probabilísticos, do estimador Kaplan-Meier, e, de acordo com o estimador e modelos propostos concluiu-se que o modelo Weibull é o que mais se ajusta a função de sobrevivência.

Atualmente os estudos de confiabilidade são cada vez mais desejáveis, pois as empresas buscam maior qualidade de seus produtos, devido à grande concorrência de mercado. Sendo assim, no trabalho aqui analisado foi possível obter uma realização de estudos de confiabilidade dos sistemas e equipamentos, através da modelagem desses sistemas, cujo as taxas de risco ou intensidade de falha variaram aleatoriamente, o que é um dos fatores mais importantes em uma política de manutenção, diminuindo os gastos com a garantia e possibilitando a extensão da mesma. E todo esse processo é um grande diferencial de mercado.

O termo confiabilidade refere-se basicamente a situações na engenharia, entretanto, condições similares ocorrem em outras áreas em que se usam as mesmas técnicas de análise de dados, como na área clínica, geralmente relacionada a dados biomédicos, também chamados de análise de sobrevivência; na área de ciências sociais, em que várias situações de interesse têm como resposta o tempo entre eventos; em estudos criminalistas, onde estudam o tempo entre a liberação de presos e a ocorrência de crimes; os estudiosos do trabalho, que se concentram em mudanças de empregos, desempregos, promoções e aposentadorias; os demógrafos, com nascimentos, mortes, casamentos, divórcios e migrações, entre outros.

4. Agradecimentos

Agradeço ao CNPq, pelo auxílio financeiro que possibilitou a realização deste trabalho, ao meu orientador, a quem devoto a mais sincera e efusiva admiração, ao CEFET/GO Inhumas e aos professores.

5. Referências Bibliográficas

BOLDRINI, J. L.; COSTA, S. I. R.; FIGUEIREDO, V. L.; WETZLER, H. G. **Álgebra Linear**. São Paulo: Harper & Row do Brasil, 1980.

BRESSAN, Graça. **Modelagem e Simulação de Sistemas Computacionais**. LARC-PCS/EPUSP, 2002.

COLOSIMO, E.A.; GIOLO, S.R. **Análise de sobrevivência aplicada**. São Paulo: Edgard Blücher, 2006.

HENLEY, R.J.; KUMAMOTO, H.. **Reliability engineering and risk assessment**. Pretience-Hall, Inc., Englewood Cliffs, 1981.

LOUZADA NETO, F.; MAZUCHELI, J.; ACHCAR, J.A.. **Análise de Sobrevivência e Confiabilidade**. São Carlos-SP, 2002.

RESTUM, G. C.. **Novas aplicações civis e militares das cadeias de Markov**. Niterói: EdUFF, 1999.

VOLLERTT JUNIOR, J. R. Confiabilidade e falhas de campo: um estudo de caso para melhoria da confiabilidade de um produto e do reparo, através de um procedimento sistemático de coleta de dados. Universidade Federal de Santa Catarina. Florianópolis-SC, 1996.

WEBER, T. S. **Um roteiro para exploração dos conceitos básicos de tolerância a falhas**. Instituto de Informática – UFRGS, 2005.